AMENDMENTS TO THE SPECIFICATION

Please amend the Specification as indicated in the replacement paragraphs below.

Following the paragraph ending at page 8, line 13, insert the following **new** paragraphs:

The terms "Zif268" and "SEQ. ID NO. 1" are used interchangeably herein.

The terms "FSD-1" and "SEQ. ID NO. 2" are used interchangeably herein.

The terms "sequence A" and "SEQ. ID NO. 3" are used interchangeably herein.

The terms "sequence B" and "SEQ. ID NO. 4" are used interchangeably herein.

The terms "Gβ1" and "SEQ. ID NO. 5" are used interchangeably herein.

The terms "sequence C" and "SEQ. ID NO. 6" are used interchangeably herein.

The terms "sequence D" and "SEQ. ID NO. 7" are used interchangeably herein.

The terms "Gβ1-c3b4" and "SEQ. ID NO. 8" are used interchangeably herein.

Please replace the paragraph at page 9, lines 3-12, with the following:

According to the method of the invention, a reduced virtual representation is first constructed for the predefined 3D structure. The reduced representation is ~~may be~~ obtained by the methodology originally developed by Herzyk and Hubbard for use with dynamic simulated annealing [Herzyk P. and Hubbard R.E. Proteins 17:310-324 (1993)]. According to this methodology, the amino acids are represented by virtual spherical atoms, wherein the main chain of the protein, polypeptide or any other suitable polymer is represented by one virtual atom per residue located at the Ca position and the side chains are represented by one or more additional virtual atoms. The number of additional virtual atoms depends on the size and chemical composition of the specific side chain.

Following the paragraph ending at page 9, line 12, insert the following **new** paragraph:

In the 'reduced representation' disclosed in the Herzyk et al reference, each amino acid residue is represented by one or more spheres, and at least one amino acid is represented by two or more spheres. The virtual bonds and excluded volume of these atoms were parameterized by analysis of 83 protein structures determined by

x-ray crystallography. The authors used this representation in restraint satisfaction calculation with dynamic simulated annealing in the context of NOE distance constraints. They compared the results obtained with this reduced representation to previous results obtained using an all-atom representation for the determination of the structure of crambin, echistatin and protein G from their experimental NOE data. They found that their 'reduced representation' permitted a 30-fold decrease in computer time for generating a single structural solution, and a 20-fold decrease in computer time to produce an acceptable structure, compared to the use of an all-atom model. The RMS distances between the reduced representation model and the all atom model were 1.5A to 1.7A.

Please replace the paragraph at page 17, line 24 to page 18, line 10, with the following:

In order to examine ~~examined~~ the method according to the invention the ββα motif typified by the zinc finger DNA binding module in the zinc finger protein, Zif268 was used. Zif268 is a well recognized protein which has the sequence shown in SEQ. ID NO. 1. This protein is small enough to be both computationally and experimentally tractable, yet large enough to form an independently folded structure in the absence of disulfide bonds or metal binding. Although this motif consists of fewer than 30 residues, it

does contain sheet, helix and turn structures. By the method and system of the invention the entire amino acid sequence: the buried core, the solvent exposed surface and the boundary between core and surface, except for the Gly27, which was not mutated during the simulation, was computed. The input coordinates are those of residues 33-60 of the native proteins obtained from the X-ray structure coordinate of Zif268 immediate early gene (krox-24) complex with an 11 base pair DNA fragment (Protein Data Bank (PDB) code: 1ZAA), as determined at 2.1Å resolution by Pavletich and Pabo [Pavletich N.P., Pabo Science 252:809-817 (1991)]. Recently, this protein was also analyzed by Dahiyat & Mayo(13).

Please replace Table 1 and its corresponding legend at page 18, lines 18-22, with the following:

Residue No	5	10	15	20	25
2nd Struct(1)	E E E T T E E E H H H H H H H H H H H H				
SEQ. ID NO. 1 ~~Zif268~~(2)	K P F Q C R I C M R N F S R S D H L T T H I R T H T G E				
SA(3)	e e i i b e e e e e e i e e e e e i e e i i e e i e e e				
D&M's SA(4)	e e i e b e i e e e e i e e e e e i e e i i e e i e e e				

(1) Secondary structure containing Extended sheet (E), Turn (T) or Helix (H);
(2) SEQ. ID NO. 1 corresponds to the Zif268 wild-type sequence written in one letter code;
(3) Solvent accessibility as determined by the invention, categorized as buried (b), exposed (e) or intermediate (i);
(4) Solvent accessibility as determined by Dahiyat and Mayo(13).

Please replace the section of the specification, including Table 2A and its corresponding legend, at page 20, line 20 to page 21, line 12, with the following:

Tables 2A and 2B present the lowest energy sequences obtained in the first and second sets (A and B respectively), aligned with the second zinc finger module of the DNA binding protein Zif268 and with D&M designed sequence, FSD-1, SEQ. ID NO. 2. The coordinates used for the FSD-1 ββα motif score evaluations are the experimental NMR coordinates (PDB code 1FSD), which were found by D&M(13). All the energy scores in Table 2B were calculated according the method of the present invention's reduced representation of amino acids and its scoring function. A and B scores were found to be lower than both Zif268 score (without considering the His2Cys2 Zn-binding interactions which are not included in the scoring function), and the FSD-1 score. The energy score of the most stable sequence, A, corresponding to SEQ. ID NO 3, is -351.8kcal/mol. This score is lower than Zif268 score by 111.3kcal/mol which is a significant difference (not taking ~~tacking~~ into account the Zn interactions). The relative stability of both SEQ. ID NO 3 and SEQ. ID NO 4 ~~A and B sequences~~ in comparison to SEQ. ID NO. 2 ~~the FSD-1 sequence~~, may be in part due to the fact SEQ. ID NO. 2 ~~the FSD-1 sequence~~ was designed with a different scoring function.

Table 2A – The most stable sequence obtained for the Zinc finger

2nd struc.[1]			E	E	E		T	T		E	E	E			H	H	H	H	H	H	H	H	H	H	H			
SA[2]			i	i	b							i					i				i	i				i			
D&M's SA[2]			i		b		i					i					i				i	i				i			
Position					5					1					1				2				2						
										0					5				0				5						
SEQ. ID NO. 2 ~~FSD-1~~[3]	Q	Q	Y	T	A	K	I	K	G	R	T	F	R	N	E	K	E	L	R	D	F	I	E	K	F	K	G	R	
SEQ. ID NO. 1 ~~Zif268~~[4]	K	P	F	Q	C	R	I	C	M	R	N	F	S	R	S	D	H	L	T	T	H	I	R	T	H	T	G	E	
SEQ. ID NO. 3 [[A]][5]	E	H	M	F	V	H	H	H	T	T	R	F	S	S	H	T	S	F	T	S	F	L	R	S	M	Q	G	R	
SEQ. ID NO. 4 [[B]][6]	Q	H	M	T	V	H	F	H	N	T	T	F	S	H	H	S	S	L	S	T	F	L	Q	S	F	Q	G	R	

[1] Secondary structure containing Extended sheet (E), Turn (T) or Helix (H);
[2] Solvent Accessibility as determined by the invention, categorized as buried (b), exposed (e) or intermediate (i) (all other positions are exposed);
[3] The sequence as designed by D&M[13];
[4] SEQ. ID NO. 1 corresponds to wild-type Zif268 sequence written in one letter code;
[5] Sequence obtained by the present invention using the SA calculation described herein;
[6] Sequence obtained by the present invention using D&M SA fitted assignments.

Please replace the section at page 22, line 21 to page 23, line 19, with the following:

3. Positions 21 and 25 of the optimal sequences were selected to be Phe or Met (position 21) and Leu (position 25) side chains. In the original SEQ. ID NO. 1 ~~Zif268~~, these positions were occupied by the zinc binding His residue. These positions are more than 80 percent buried. Position 5, which is 100 percent buried, was predominantly selected to be Val. The other boundary positions

demonstrate the steric constrains on buried residues by packing similar side chains to those of the original SEQ. ID NO. 1 ~~Zif268 sequence~~.

4. In the helix region (residues 15-26) the algorithm placed two Leu side chains and one Gln, which are good helix forming residues, in SEQ. ID NO. 4 ~~sequence B~~, and one Leu and one Gln in SEQ. ID NO. 3 ~~sequence A~~.

5. In both SEQ. ID NO. 3 and SEQ. ID NO. 4 ~~A and B sequences~~, position 5 on the exposed sheet surface was selected by the algorithm to be Val, which is a very good β-sheet forming residue, and positions 4 and 10 (and 11 only in SEQ. ID NO.4 ~~sequence B~~) were selected to be Thr, which is also a good β-sheet forming residue.

6. Alignment of the optimal stable SEQ. ID NO. 4 ~~sequence (B)~~ and SEQ. ID NO. 1 ~~Zif268~~ indicates that 4 out of 27 residues (not including residue 27 that remains Gly throughout the simulation) are identical (15%) and 11 are similar (including the identical 40.7%). D&M obtained similar values, with 5 identical residues (18.5%) and 12 similar (44.4%).

7. Alignment of the SEQ. ID NO. 4 ~~sequence B~~ and SEQ. ID NO. 2 ~~FSD-1~~ indicates that 5 out of 27 residues are identical between the sequences (18.5%) and 11 are similar (including identical 40.7%).

7.1 Secondary structure prediction of the designed sequence SEQ. ID NO. 3 and SEQ. ID NO. 4 ~~Sequence A and B~~ were further examined by secondary structure prediction by the SSPAL predictor at Sanger Centere [Salamov A. A. and Solovyev V. V. J. Mol. Biol. 247:11-15 (1995)], which enable to predict the secondary structure of a protein according to its primary structure (amino acid sequence). By these programs both SEQ. ID NO. 3 and SEQ. ID NO. 4 ~~A and B sequences~~ were predicted to have the desired Zinc finger motif. Table 3 presents the secondary structure of the native protein (Zif268) according to the Protein Data Bank (PDB), and SEQ. ID NO. 3 and SEQ. ID NO. 4 ~~A and B~~ secondary structure prediction, according to SSPAL algorithm at Sanger Centre. SEQ. ID NO. 3 ~~A~~ was predicted to have one α-helix (designated H) and two β-strands (designated E)(the ββα motif) while the predicted secondary structure to SEQ. ID NO. 4 ~~B~~ contained only one α-helix and one β-strand.

Please replace Table 3 and its corresponding legend at page 24, lines 3-5, with the following:

Table 3 – Secondary structure of predicted primary structures *A* and *B*.

Position	5	1 0	1 5	2 0	2 5
SS[(1)] PDB	E E E T T	E E E	H H H H H H H H H H H		
SEQ. ID NO. 1 ~~Zif268~~	K P F Q C R I C M R N F S R S D H L T T H I R T H T G E				
SS[(1)] *A*	E E E E	E E	H H H H H H H H H H H		
SEQ. ID NO. 3 [[A]]	E H M F V H H H T T R F S S H T S F T S F L R S M Q G R				
SS[(1)] *B*	E E E E E E		H H H H H H H H		
SEQ. ID NO. 4 [[B]]	Q H M T V H F H N T T F S H H S S L S T F L Q S F Q G R				

[(1)] Secondary Structure

Please replace the paragraph at page 24, line 7 to page 25, line 3, with the following:

The reduced representation of the lowest energy designed SEQ. ID NO. 3 and SEQ. ID NO. 4 ~~sequences A and B~~, was expanded to an all-atom representation, using the molecular mechanics package CHARMM. The input for this experiment was the backbone coordinates of the native protein, the new designed residues and the dihedral angles of each position along the designed sequence derived from the rotamer with the lowest energy score. The number of atoms of SEQ. ID NO. 3 and SEQ. ID NO. 4 ~~A and B~~ after expansion to all atoms, were 459 and 446, respectively. Energy minimization was performed for SEQ. ID NO. 3 and SEQ. ID NO. 4 ~~A and B's~~ side chains as well as to SEQ. ID NO. 1 ~~Zif268~~ side chains using CHARMM

forcefield, the SHAKE algorithm [Van Gunsteren W.F. & Berendsen H.J.C. Mol. Phys. 34:1311 (1977)], a dielectric constant of ε=1 and a 12Å energy cutoff. The minimization included 200 steps of SD (Steepest Descent) and then additional 500 steps of ABNR (Adopted Basis Newton Raphson). After minimization, each of the three structures were embedded in an 18Å water sphere which included ~1870 water molecules of type TIP3P [Jorgenes W.L. et al. J Chem. Phys. 79:926-935 (1983)]. Each of the water-protein systems of SEQ. ID NO. 3 and SEQ. ID NO. 4 ~~A and B~~ were simulated for 500ps at 300K (with a 16Å energy cutoff) and a sample of 2000 conformations was collected from the resulting molecular dynamics trajectory. It was found that the secondary structure of the proteins was maintained during the molecular dynamic simulations.

Please replace Table 4, at page 28, lines 24 1-5, with the following:

2nd structure(1)	**E**	**E**		**E**	**H**	**H**		**E**	
Solvent accessibility	b	b	i	i	i	i	b	i	
Position	3	7	16	18	25	29	39	43	**Energy score(2) kcal/mol**
SEQ. ID NO. 5 ~~Gb1~~	Y	L	T	T	T	V	V	W	-676.4
SEQ. ID NO. 6 [[C]]	L	L	L	F	L	L	V	M	-778.6
SEQ. ID NO. 7 [[D]]	L	L	F	L	L	S	V	F	-788.8
SEQ. ID NO. 8 ~~Gβ1-c3b4~~(3)	F	I	I	I	Q	I	I	I	

Please replace the paragraph at page 28, lines 12-17, with the following:

2. In most sequences Thr25 was mutated to Leu, which has a better helix propensity. Examining of SEQ. ID NO. 6 and SEQ. ID NO. 7 ~~sequences C and D~~ by SSPAL predictor at Sanger Centre (as described hereinbefore) and by the PHD predictor at EMBL [Rost B and Sander C. J. Mol. Biol. 232:584-599 (1993)] showed that mutations T25L and V29L maintained the secondary structure of the α-helix.

Please replace the paragraph at page 29, line 11 to page 30, line 7, with the following:

1.4 Minimization of the redesigned sequences in comparison to native Gβ1 of SEQ. ID NO. 5

The reduced representation of the designed SEQ. ID NO. 6 ~~sequence C~~, which was the sequence with the lowest energy score in the first set of simulations, where the conformation of 48 non-mutated residues was kept fixed, was expanded to its all-atom representation, using CHARMM [Brooks B.R. et al. (1983) ibid.], in the same manner as described hereinbefore. In general, the information provided as input was the backbone coordinates of the native protein, the new residues and the rotamer dihedral angles at each position along the chain. The number of atoms in SEQ. ID NO.

6 ~~sequence C~~ was 865. Energy minimization was performed for the side chains of this sequence as well as for Gβ1's side chains using CHARMM force-field [Mackerel A.D. et. Al. (1998) ibid.], a dielectric constant ε=1 and a 16Å energy cutoff. The minimization included 800 steps of SD and then additional 1100 steps of ABNR. The average energies of the sequences after minimization were:-

SEQ. ID NO. 5 ~~Gβ1~~: -804.2 kcal/mol

SEQ. ID NO. 6 [[C]]: -822.7 kcal/mol

The above results suggest that the mutations obtained by the methodology disclosed herein are tolerable which may lead to the designed of a more stable protein.

The difference between the energies of the native sequence SEQ. ID NO. 5 ~~Gβ1~~ and SEQ. ID NO. 6 ~~sequence C~~, based on the present invention's scoring function and on CHARMM's force-field (after dynamics) were 7% and 16% respectively, which strengthens the conclusion that the method and system of the present invention provide a reliable tool for designing de novo proteins.

The above statement is further strengthened in light of Figure 7, which show a comparison of the 3D structure of SEQ. ID NO. 5 ~~Gβ1~~ and SEQ. ID NO. 6 ~~sequence C~~.